

(19) World Intellectual Property Organization  
International Bureau



PCT

(43) International Publication Date  
12 September 2008 (12.09.2008)

(10) International Publication Number  
**WO 2008/107002 A1**

(51) International Patent Classification:  
**G06K 9/00** (2006.01)

(21) International Application Number:  
**PCT/EP2007/006540**

(22) International Filing Date: 23 July 2007 (23.07.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/892,883 5 March 2007 (05.03.2007) US

(71) Applicant (for all designated States except US): **FOTONATION VISION LIMITED** [IE/IE]; Galway Business Park, Dangan, Galway (IE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **STEINBERG, Eran** [IL/US]; 137 Granville Way, San Francisco, CA 94127 (US). **BIGIOI, Petronel** [RO/RO]; 57 Sceilg Ard, Headford Road, County Galway (IE). **CORCORAN, Peter** [IE/IE]; Cregg, Claregalway, Galway (IE). **GANGEA, Mihnea** [RO/RO]; Sos. Nicolae Titulescu Nr.

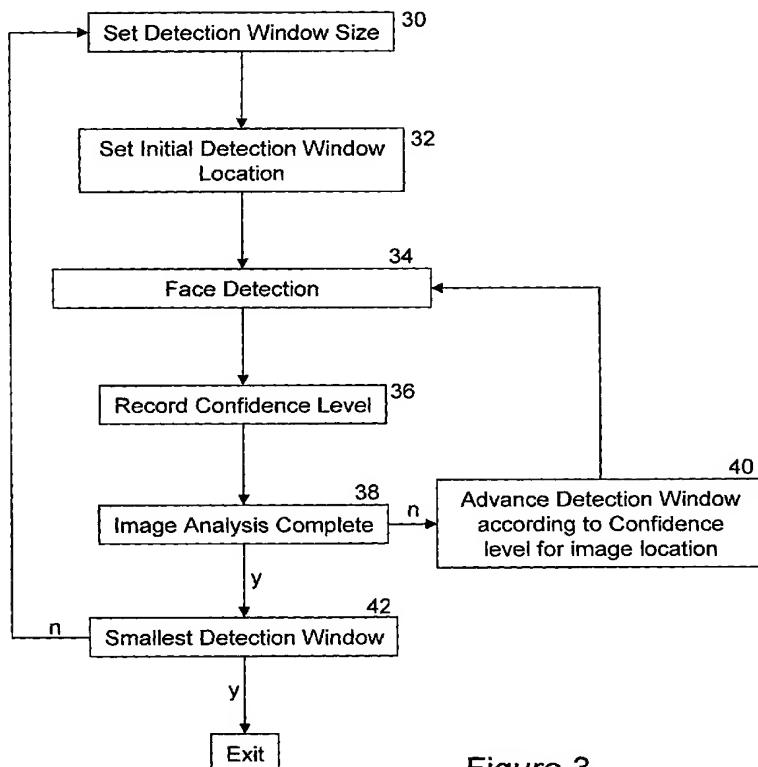
10, Bloc 20, Scara A et 9, Ap. 44, Sector 1, Bucuresti (RO). **PETRESCU, Stefan, Mirel** [RO/RO]; Bl.73, sc.3, et.2, ap.37, Str. Cap. Gheorghe Ion, Nr.4, Sector 4, Mun.Bucuresti (RO). **VASILIU, Andrei** [RO/RO]; Str. Armenis Nr. 8-10. Bl. J5, Apt. 84, Sector 3, Bucuresti (RO). **COSTACHE, Gabriel** [RO/IE]; 19 Lui Na Greine, Knocknacarra, Galway (IE). **DRIMBAREAN, Alexandru** [RO/IE]; 60 Garraun Ard, Doughiska, Galway (IE).

(74) Agents: **CONOR, Boyce** et al.; F.R. Kelly & Co., 27 Clyde Road, Ballsbridge, Dublin 4 (IE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: FACE SEARCHING AND DETECTION IN A DIGITAL IMAGE ACQUISITION DEVICE



(57) Abstract: A method of detecting a face in an image includes performing face detection within a first window of the image at a first location. A confidence level is obtained from the face detection indicating a probability of the image including a face at or in the vicinity of the first location. Face detection is then performed within a second window at a second location, wherein the second location is determined based on the confidence level.

Figure 3



- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL,

PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *with international search report*

## FACE SEARCHING AND DETECTION IN A DIGITAL IMAGE ACQUISITION DEVICE

## FIELD OF THE INVENTION

5       The present invention provides an improved method and apparatus for image processing in digital image acquisition devices. In particular the invention provides improved performance and accuracy of face searching and detection in a digital image acquisition device.

10

## BACKGROUND OF THE INVENTION.

Several applications such as US published application no. 2002/0102024 to inventors Jones and Viola relate to fast-face detection in digital images and describe certain algorithms. Jones and Viola describe an algorithm that is based on a cascade of increasingly refined 15 rectangular classifiers that are applied to a detection window within an acquired image. Generally, if all classifiers are satisfied, a face is deemed to have been detected, whereas as soon as one classifier fails, the window is determined not to contain a face.

An alternative technique for face detection is described by Froba, B., Ernst, A., "Face detection with the modified census transform", in Proceedings of 6<sup>th</sup> IEEE Intl. Conf. on 20 Automatic Face and Gesture Recognition, 17-19 May 2004 Page(s): 91 – 96. Although this is similar to Violla-Jones each of the classifiers in a cascade generates a cumulative probability and faces are not rejected if they fail a single stage of the classifier. We remark that there are advantages in combining both types of classifier (i.e. *Violla-Jones* and *modified census*) within a single cascaded detector.

25       Figure 1 illustrates what is described by Jones and Viola. For an analysis of an acquired image 12, the detection window 10 is shifted incrementally by dx pixels across and dy pixels down until the entire image has been searched for faces 14. The rows of dots 16 (not all shown) represent the position of the top-left corner of the detection window 10 at 30 each face detection position. At each of these positions, the classifier chain is applied to detect the presence of a face.

Referring to Figures 2a and 2b, as well as investigating the current position, neighboring positions can also be examined, by performing small oscillations around the current detection window and/or varying slightly a scale of the detection window. Such

oscillations may vary in degree and in size creating consecutive windows with some degree of overlap between an original window and a second window. The variation may also be in the size of the second window.

5 A search may be performed in a linear fashion with the dx,dy increments being a pre-determined function of image resolution and detection window size. Thus, the detection window may be moved across the image with constant increments in x and y directions.

A problem with linear searching occurs when the window size decreases, such as when attempting to detect small faces, and the number of sliding windows that are to be analyzed increases quadratically to the reduction in window size.. This results in a 10 compounded slow execution time, making “fast” face detection otherwise unsuitable for real-time embedded implementations.

US Application No. 11/464,083, filed August 11, 2006, which is assigned to the same assignee as the present application, discloses improvements to algorithms such as those described by Jones and Viola, and in particular in generating a precise resolution 15 corresponding to a representation of an image, such as an integral image or a Gaussian image, for subsequent face detection.

#### SUMMARY OF THE INVENTION

20 A method of detecting a face in an image includes performing face detection within a first window of the image at a first location. A confidence level is obtained from the face detection indicating a probability of the image including a face at or in the vicinity of the first location. Face detection is performed within a second window at a second location that is determined based on the confidence level.

25

A number of windows that are analyzed is advantageously significantly reduced for a same face detection quality, and so faster face searching is provided, even in the case of small faces, therefore allowing acceptable performance for face detection in real-time embedded implementations such as in digital cameras, mobile phones, digital video cameras and hand-held computers.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments will now be described, by way of example, with reference to the accompanying drawings, in which:

Figure 1 illustrates schematically an image being processed by a conventional face detection process;

5       Figure 2(a) illustrates a detection window oscillating diagonally around an initial position;

Figure 2(b) illustrates a smaller scale detection window oscillating transversely around the initial position;

10      Figure 3 is a flow diagram of a method of face searching and detection according to a preferred embodiment;

Figure 4 illustrates schematically an image being processed according to a preferred embodiment; and

15      Figure 5 is a flow diagram illustrating post-processing of a detected face region prior to face recognition.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

An improved method of face searching and detection in a digital image acquisition device is described that calculates x and/or y increments of a detection window in an adaptive fashion.

In face detection processes, during analysis of a detection window and/or while oscillating around the detection window, a confidence level can be accumulated providing a probabilistic measure of a face being present at the location of the detection window. When the confidence level reaches a preset threshold for a detection window, a face is confirmed 25 for location of the detection window.

Where a face detection process generates such a confidence level for a given location of detection window, in a preferred embodiment, the confidence level is captured and stored as an indicator of the probability of a face existing at the given location. Such probability may reflect confidence that a face has been detected, or confidence that there is no face 30 detected in the window.

Alternatively, where a face detection process applies a sequence of tests each of which produce a Boolean "Face" or "No face" result, the extent to which the face detection process has progressed through the sequence before deciding that no face exists at the

location can be taken as equivalent to a confidence level and indicating the probability of a face existing at the given location. For example, where a cascade of classifiers fails to detect a face at a window location at classifier 20 of 32, it could be taken that this location is more likely to include a face (possibly at a different scale or shifted slightly) than where a cascade of classifiers failed to detect a face at a window location at classifier 10 of 32.

Referring now to Figure 3, face searching and detection according to one embodiment, begins by selecting the largest size of detection window at step 30 and positioning the window at the top left corner of an image at step 32.

Alternatively, if particular regions of an image have been identified through some pre-processing as being more likely to include a face, the detection window can be located at a suitable corner of one such region and the embodiment can be applied to each such region of the image in turn or in parallel. Examples of such pre-processing include identifying regions of the image which include skin as being candidate face regions.

In this regard, it is possible to create a skin map for an acquired image where the value of a pixel within the skin map is determined by its probability of being a skin pixel.

There are many possible techniques for providing a skin map, for example:

(i) "Comparison of Five Color Models in Skin Pixel Classification", Zarit et al presented at ICCV '99 International Workshop of Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, contains many references to tests for skin pixels;

(ii) U.S. Pat. No. 4,203,671, Takahashi et al., discloses a method of detecting skin color in an image by identifying pixels falling into an ellipsoid in red, green, blue color space or within an ellipse in two dimensional color space;

(iii) US Pat. No. 7,103,215 describes a method of detecting pornographic images, wherein a color reference database is prepared in LAB color space defining a plurality of colors representing relevant portions of a human body. A questionable image is selected, and sampled pixels are compared with the color reference database. Areas having a matching pixel are subjected to a texture analysis to determine if the pixel is an isolated color or if other comparable pixels surround it; a condition indicating possible skin;

(iv) US 11/624,683 filed January 18, 2007 (Ref: FN185) discloses real-valued skin tests for images in RGB and YCC formats. So, for example, where image information is available in RGB format, the probability of a pixel being skin is a function of the degree to which L exceeds 240, where  $L=0.3*R+0.59*G+0.11*B$ , and/or the degree to which R exceeds G + K and R exceeds B + K where K is a function of image saturation. In YCC format, the

probability of a pixel being skin is a function of the degree to which Y exceeds 240, and/or the degree to which Cr exceeds  $148.8162 - 0.1626*Cb + 0.4726*K$  and Cr exceeds  $1.2639*Cb - 33.7803 + 0.7133*K$ , where K is a function of image saturation.

It will also be understood that many different techniques exist to provide a binary skin/not-skin classification (typically based on simple thresholding). So, it can be understood that some pixels may qualify as skin under one binary technique and as not-skin under a second technique. So in alternative implementations, several binary techniques can be combined, so that pixels may be ranked according to a number of criteria to obtain a relative probability that any particular pixel is skin. It is advantageous to weight different skin detection techniques according to image capture conditions, or according to data analyzed from previous image frames.

Where multiple skin classification techniques are implemented in a parallel hardware architecture it becomes possible to combine the outputs of multiple skin classification techniques in a single analysis step, quickly generating a refined skin probability for each image pixel as it becomes available from the imaging sensor. In one preferred embodiment this refined skin probability is represented as a grayscale value,  $2^N$  where  $N > 1$  ( $N=1$  represents a simple binary mask of skin/not-skin). In any case, once an image pixel is classified by a non-binary algorithm it may be considered as a grayscale representation of skin probability.

In assessing whether various sizes and locations of windows in an image might include portions of skin, it can be advantageous to use the integral image techniques disclosed in US 2002/0102024, Viola-Jones with the skin map probability values produced for an image.

In such an integral image, each element is calculated as the sum of intensities i.e. skin probabilities of all points above and to the left of the point in the image. The total intensity of any sub-window in an image can then be derived by subtracting the integral image value for the top left point of the sub-window from the integral image value for the bottom right point of the sub-window. Also intensities for adjacent sub-windows can be efficiently compared using particular combinations of integral image values from points of the sub-windows.

Thus the techniques employed to construct an integral image for determining the luminance of a rectangular portion of the final image may equally be employed to create a skin probability integral image. Once this integral image skin map (IISM) is created, it enables the skin probability of any rectangular area within the image to be quickly

determined by simple arithmetic operations involving the four corner points of the rectangle, rather than having to average skin values over the full rectangle.

In the context of a fast face detector as described in the remainder of this specification, it can be understood that obtaining a rapid calculation of the averaged local skin pixel probability within a sub-window enables the skin probability to be advantageously employed either to confirm a local face region, or to be used as an additional, color sensitive classifier to supplement conventional luminance based Haar or census classifiers.

Alternatively or in combination with detection of skin regions, where the acquired image is one of a stream of images being analyzed, the candidate face regions might be face regions detected in previous frames, such as may be disclosed at US Application No.

10 11/464,083, (Ref: FN143) filed August 11, 2006.

Figure 2a illustrates the detection window oscillating diagonally around an initial position (outlined in bold). Figure 2b illustrates a smaller scale detection window oscillating transversely around the initial position before further face detection is performed. These 15 oscillations  $\Delta x, \Delta y$  and scale changes  $\Delta s$  are typically smaller than the  $\Delta x, \Delta y$  step of the detection window. A decision as to scale of oscillation depends on results of applying the search algorithm on the initial window. Typically, a range of about 10-12 different sizes of detection window may be used to cover the possible face sizes in an XGA size image.

Returning to the operation of the main face detector, we note that face detection is 20 applied for the detection window at step 34, and this returns a confidence level for the detection window. The particular manner in which the detection window oscillates around a particular location and the calculation of the confidence level in the preferred embodiment is as follows:

Once a given detection window location has been tested for the presence of a face, the 25 window is sequentially shifted by  $-\Delta x, -\Delta y$ ;  $+\Delta x, -\Delta y$ ;  $+\Delta x, +\Delta y$ ; and  $-\Delta x, +\Delta y$  (as shown in Figure 2(a)) and tested at each of these four locations. The confidence level for the window location and four shifted locations is accumulated. The confidence level may then be ported to each new window based on the new window size and location. If a target face-validation confidence threshold is not reached, the detection window is shrunk (indicated by 30  $\Delta s$ ). This smaller detection window is tested, then sequentially shifted by  $-\Delta x, 0$ ;  $+\Delta x, 0$ ;  $0, +\Delta y$ ; and  $0, -\Delta y$  (as shown in Figure 2(b)) and tested at each of these four locations. The confidence level for these five locations of the smaller scale detection window is added to the previous confidence level from the larger scale window.

The confidence level for the detection window location is recorded at step 36.

If the detection window has not traversed the entire image/region to be searched at step 38, it is advanced as a function of the confidence level stored for the location at step 40.

In the preferred embodiment, where the confidence level for an immediately previous 5 detection window at the present window size has exceeded a threshold, then the x and y increment for the detection window is decreased.

Referring now to Figure 4, which shows how in the preferred embodiment, the shift step is adjusted when the confidence level for a location signals the probability of a face in the vicinity of the location. For the first four rows of searching, a relatively large increment 10 is employed in both x and y directions for the detection window. For the location of detection window 10(a), however, it is more than likely that the oscillation of the window in the bottom-right direction will provide the required confidence level of the face 14 being at the location. As such, the increment for the detection window in the x and y directions is decreased. In the example, the confidence level will remain above the determined threshold 15 until the detection window location passes to the right of the line t12. At this time, the x increment reverts to the original large increment. Having incremented by the small increment in the y direction, the detection window is advanced on the next row with a large x increment until it reaches the line t11. Either because the confidence level for this location will again exceed the required threshold or indeed because it did for the previous row, the x increment is 20 again decreased until again the detection window passes to the right of the line t12. This process continues until the detection window arrives at location 10(b). Here, not alone is the confidence level for increased resolution face detection reached, but the face 14 is detected. In the preferred embodiment, this causes both the x and y increments to revert to original large increments.

25 If a face is not detected in a region following a confidence level triggering at a face-like (but not an actual face) position, the x and y increments return to their original relaxed value, when over the whole extent of a row, the confidence levels do not rise above the threshold level. So for example, in the row after the detection window passes location 10(c), no detection window will produce a confidence level above the threshold and so after this 30 row, the y increment would revert to its relaxed level, even if a face had not been detected at location 10(b).

Once the image and/or its regions have been traversed by a detection window of a given size, unless this has been the smallest detection window at step 42 of Figure 3, the next smallest detection window is chosen at step 30, and the image traversed again.

In certain embodiments, when the confidence level for an immediately previous 5 detection window at the present window size exceeds a threshold, a change in dx,dy for a detection window is triggered. However, this change could equally and/or additionally be a function of or be triggered by the confidence level for a bigger detection window or windows at or around the same location.

In certain embodiments, detection windows are applied from the largest to the 10 smallest size and so it is assumed that a given location has been checked by a larger sized detection window before a given sized detection window, so indicating that if a face has not been detected for the larger sized detection window, it is to be found near that location with a smaller sized detection window. Alternatively, it can indicate that even if a face has been 15 found at a given location for a larger sized detection window, there is a chance that the face might be more accurately bounded by a smaller sized detection window around that location when subsequently applied.

As many more windows may be employed when looking for smaller size faces than 20 larger faces, where confidence levels from larger detection windows are used to drive the increments for smaller detection windows, the savings made possible by embodiments of the present invention are greater than if smaller detection windows were applied first.

In the embodiments described above, for a given detection window size, either a large or small x or y increment is employed depending on whether or not a face is likely to be in the vicinity of a detection window location. However, the increment can be varied in any suitable way. So for example, the increment could be made inversely proportional to the 25 confidence level of previous detection windows applied in the region.

Alternatively, instead of returning a quasi-continuous value as described above, the confidence level returned by the face detection process 34 could be discretely-valued indicating either: (i) no face; (ii) possible face; or (iii) face, each causing the advance step 40 to act as set out in relation to Figure 4.

The detection window does not have to move along a row. Instead, its progress in 30 each of the x and y directions may be adjusted from one increment to the next as a function of the confidence level of previous detection windows applied in the region.

The embodiments described above can be implemented in a digital image processing device such a digital stills camera, a digital video camera, camera phone or the like. The embodiments due to their computational efficiency can be implemented within a real-time face detection function which for example enables the device to highlight with a respective boundary (corresponding to a detection window) in a viewfinder faces detected in an acquired image or image stream.

Alternatively or in addition, the embodiments can be implemented within an off-line face detection function either within a digital image processing device or in a connected computing device to which an image is transferred or which has access to the image, to provide more efficient face detection.

Alternatively or in addition, the detected face regions can be employed with image post-processing functions such as red-eye detection and/or correction, or for example face expression detection and/or correction, or face recognition.

Where the detected face regions are employed in facial recognition, as many facial recognition systems remain sensitive to slight variations in either facial rotation or size, it is advantageous to apply post-processing measures in order to optimize the accuracy of facial recognition. This is because, even where frontal face regions are detected and saved, these regions may not be optimally aligned or scaled for the purposes of face recognition. Also, it should be noted that many images captured are consumer images and that subjects in such images will rarely maintain their faces in a squarely facing frontal position at the time of image acquisition.

Where as in the embodiment above, the face detection employed is highly optimized for speed and for the accurate determination of the presence of a face region, face detection is typically not optimized to accurately match the precise scale, rotation or pose of a detected face region.

There are many techniques known in the prior art for achieving such normalization, however, in an embedded imaging device, such as a digital camera, where processing must be both compact in terms of code footprint and efficient resource usage, it can be impractical to deploy more of such complex processing.

Thus, in one embodiment the face detector, already available within the image acquisition device, can be re-purposed for use in post-processing of detected/tracked face regions. In the embodiment, a supplementary *frontal face* detector which is generally identical to a standard detector, but highly optimized at the training stage to detect only

frontal faces is employed. So for example, the frontal face detector would not be suitable for normal face detection/tracking where a more relaxed detector, hereafter referred to as a *standard* detector is required.

Referring now to Figure 5, in this embodiment, if a face region to which face recognition is to be applied is originally detected, step 50, with an initial probability less than a 1<sup>st</sup> threshold, the region is expanded by say, X=20% to include a surrounding peripheral region and extracted from the acquired image, step 52. This larger region is typically sufficient to contain the entire face.

A standard detector is next applied to the expanded region, step 54, but across a smaller range of maximum and minimum scales, and at finer granular resolution than would be employed across a full image.

As an example, at step 54, the detector might scale from 1.1 to 0.9 times the size of the face region determined by the original detection process, step 50, but in increments of 0.025; thus 0.9, 0.925, 0.95, 0.975, 1.00, and so on, and similarly with step size. The goal is to determine a sub-window optimized in scale and alignment within the extracted, expanded face region where the face probability is highest. Ideally, such a sub-window will exceed a 2<sup>nd</sup> threshold probability for face detection no less than the 1<sup>st</sup> threshold. If not, and if rotation is not to be applied in an attempt to improved this probability, then this face region is marked as “unreliable” for recognition, step 56.

Where the first or second thresholds are exceeded then either the sub-window for the originally detected face region or the optimized window from step 54 are expanded by say Y = 10% < X, step 58.

The *frontal face* detector is then applied to the expanded region, step 60. If a sub-window with a face detection probability above a third threshold (higher than each of the first and second thresholds is identified), step 62, then that sub-window is marked as “reliable” and is passed on to a recognition process, step 64.

Where the *frontal face* detection step fails at step 62, but we know there is a high probability face region, then it is likely that one or both of a small rotational or pose normalization is also required to produce a face region suitable for face recognition.

In one embodiment, the original X% expanded face region is next rotated through one of a number of angular displacements, say -0.2, -0.15, -0.1, -0.05, 0.0, +0.05, +0.1, +0.15 and +0.2 radians, step 66, and the fine grained standard face detection and possibly frontal face detection steps are re-applied as before.

Ideally, the face probability will increase above the required 3<sup>rd</sup> threshold as these angular rotations are applied to the extracted face region and the face region can be marked as “reliable”. It will also be seen that the potentially changing probabilities from face region rotation can also be used to guide the direction of rotation of the region. For example, if a 5 rotation of -0.05 radians increases the face detection probability but not sufficiently, then the next rotation chosen would be -0.1 radians. Whereas if a rotation of -0.05 radians decreases the face detection probability, then the next rotation chosen would be 0.05 radians and if this did not increase the face detection probability, then the face region could be marked as “unreliable” for recognition, step 56

10 As an alternative or in addition to this in-plane rotation of the face region, an AAM (Active Appearance Model) or equivalent module can be applied to the detected face region in an attempt to provide the required pose normalization to make the face region suitable for face recognition. AAM modules are well known and a suitable module for the present embodiment is disclosed in “Fast and Reliable Active Appearance Model Search for 3-D 15 Face Tracking”, F Dornaika and J Ahlberg, IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, Vol 34, No. 4, pg 1838-1853, August 2004, although other models based on the original paper by TF Cootes et al “Active Appearance Models” Proc. European Conf. Computer Vision, 1998, pp 484-498 could also be employed.

In this embodiment, the AAM model has two parameters trained for horizontal and 20 vertical pose adjustments, and the AAM model should converge to the face within the detected face region indicating the approximate horizontal and vertical pose of the face. The face region may then be adjusted by superimposing the equivalent AAM model to provide a “straightened” face region rotated out of the plane of the image, step 68.

Again, fine grained standard face detection and frontal face detection steps are re- 25 applied, and if the threshold for the detected face region(s) is not above the required probability, then small incremental adjustments of the horizontal and vertical pose may be stepped through as before until either the *frontal face* detector probability increases sufficiently to mark the face region as “reliable” or the face region is confirmed to be “unreliable” to use for face recognition purposes.

30 US Patent Application 11/752,925 filed May 24, 2007 (Ref: FN172) describes capturing face regions from a preview stream and subsequently aligning and combining these images using super-resolution techniques in order to provide a repair template for portions of a facial region in a main acquired image. These techniques may be advantageously employed,

in addition to or as an alternative to the steps above, independently or as part of a post-processing step on a face region in order to bring the face region into a substantially frontal alignment before face recognition.

In other alternative applications for detected face regions, the selected regions may be  
5 consecutively applied to a series of images such as preview images, post-view images or a video stream of full- or reduced-resolution images, or combinations thereof, where the confidence level as well as the window locations are passed from one preview image, post-view image, etc., to the next.

While an exemplary drawings and specific embodiments of the present invention have  
10 been described and illustrated, it is to be understood that the scope of the present invention is not to be limited to the particular embodiments discussed.

In addition, in methods that may be performed according to preferred embodiments herein and that may have been described above, the operations have been described in selected typographical sequences. However, the sequences have been selected and so ordered  
15 for typographical convenience and are not intended to imply any particular order for performing the operations, except for those where a particular order may be expressly set forth or where those of ordinary skill in the art may deem a particular order to be necessary.

**Claims:**

1. A method of detecting a face in an image comprising:

(a) performing face detection within a first window of said image at a first location;

5 (b) obtaining from said face detection a confidence level indicating a probability of said image including a face at or in the vicinity of said first location; and

(c) performing said face detection within a second window at a second location wherein said second location is determined based on said confidence level.

10 2. A method as claimed in claim 1, wherein direction and magnitude of displacement within said image from said first location to said second location comprise a function of said confidence level.

15 3. A method as claimed in claim 1, further comprising repeating (a) to (c) for one or more additional windows in different locations until said performing face detection results in positive detection of a face.

20 4. A method as claimed in claim 1 further comprising repeating (a) to (c) for one or more additional windows in different locations until face detection has been performed over an entire region of interest.

5. A method as claimed in claim 4, wherein the dimensions of said one or more additional windows depend on said confidence level.

25 6. A method as claimed in claim 1, further comprising repeating (a) to (c) until face detection for multiple predetermined sizes of windows has been performed over an entire region of interest of said image.

30 7. A method as claimed in claim 1, further comprising determining a set of regions of interest for said image, and repeating steps (a) to (c ) for all said regions of interest.

8. A method as claimed in claim 1, further comprising identifying at least one region of said image likely to contain a face; and repeating (a) to (c) until face detection is performed over substantially an entirety of said at least one region.

5 9. A method as claimed in claim 1, further comprising identifying at least one region of said image likely to contain a face; and repeating (a) to (c) until face detection is performed concentrically inside said at least one region.

10 10. A method as claimed in claim 1, wherein said face detection comprises applying a chain of classifiers to said windows of said image and wherein said confidence level comprises a function of a number of classifiers successfully applied to said windows.

15 11. A method as claimed in 1, wherein said face detection returns a confidence level indicating that no face is present in the vicinity of said first location, that a face may be present in the vicinity of said first location, or that a face is present at said first location.

12. A method as claimed in claim 1, wherein said second location is determined by advancing said window in an x-amount in a first direction and in an orthogonal second direction by y-amount from the first location, wherein any overlap of the first and second windows is based on a said confidence level.

20 13. A method as claimed in claim 12, wherein a size of the second window relative to a size of said first window is also based on said confidence level.

25 14. A method as in claim 12, wherein the x-amount or y-amount, or both, are less for a higher confidence level that the image includes a face in the vicinity of the first location.

30 15. A method as claimed in claim 12, wherein the x-amount or the y-amount, or both, are greater for a higher confidence level that the image includes a face at the first location.

16. A method as claimed in claim 12, wherein the x-amount or the y-amount, or both, are greater for a higher confidence level that the image does not include a face at the first location.

5 17. A method as claimed in claim 12, wherein the x-amount and the y-amount depend separately on face detection confidence levels in the first and second directions, respectively.

10 18. A method as claimed in claim 1, wherein when there is a confidence level above a threshold value that a face is detected in the vicinity of said first location, then the second location is selected such that the second window overlaps said first window to center on said face, and when there is a confidence level below said threshold, the second location is selected so that the second window does not overlap the first window.

15 19. A method as claimed in claim 1, further comprising pre-determining one or more regions of interest within the image each as having an enhanced likelihood of including a face, and locating a detection window at a suitable corner of each such region of interest, and applying (a) to (c) to each such region of interest.

20 20. A method as claimed in claim 19, wherein (a) to (c) are applied to two or more regions of interest in time periods with at least some temporal overlap.

21. A method as claimed in claim 19, wherein said one or more regions of interest comprise one or more regions including a number of skin pixels.

25

22. A method as claimed in claim 19, wherein said image is an image in a stream of images and wherein said regions of interest comprise one or more regions in which a face has been detected in a previous image of said stream.

30

23. A method as claimed in claim 1 wherein said step (c) comprises:  
responsive to said confidence level indicating a face at or in the vicinity of said first location, performing detection of a frontally aligned face within a second window at a second

location wherein said second location is determined based on a confidence level indicating a probability of a candidate region including said face at or in the vicinity of said first location.

24. A method as claimed in claim 23 comprising:

5 (d) responsive to detection of a frontally aligned face in any candidate region, selectively applying face recognition to said candidate region.

25. A method as claimed in claim 24 wherein face recognition is applied in response to any candidate region having a probability of including a face greater than a threshold.

10

26. A method as claimed in claim 23 wherein said performing detection of a frontally aligned face is performed in a candidate region including said face detected at or in the vicinity of said first location.

15

27. A method as claimed in claim 26 wherein performing detection of a frontally aligned face is performed in response to said candidate region having a probability of including a face greater than a first threshold.

28. A method as claimed in claim 26 comprising:

20

responsive to said first location having a probability of including a face less than a first threshold, extracting an extended region including said first location from said image and performing face detection in said extended region with a relatively fine granularity to provide said candidate region.

25

29. A method as claimed in claim 26 comprising:

responsive to a candidate region having a probability of including a face less than a second threshold, rotating an extended region including said face region through one of a sequence of angles and performing face detection in said rotated extended region with a relatively fine granularity to provide another candidate region.

30

30. A method as claimed in claim 26 comprising:

responsive to a candidate region having a probability of including a face less than a second threshold, rotating said face within said extended region through one of a sequence of

angles and performing face detection in said extended region with a relatively fine granularity to provide another candidate region.

31. A digital image processing device adapted to:

- 5 (a) perform face detection within a first window of an acquired image at a first location;
- (b) obtain from said face detection a confidence level indicating a probability of said image including a face at or in the vicinity of said first location; and
- 10 (c) perform said face detection within a second window at a second location wherein said second location is determined based on said confidence level.

32. A computer program product which when executed in a digital image processing device is operable to perform the steps of:

- 15 (a) performing face detection within a first window of an acquired image at a first location;
- (b) obtaining from said face detection a confidence level indicating a probability of said image including a face at or in the vicinity of said first location; and
- (c) performing said face detection within a second window at a second location wherein said second location is determined based on said confidence level.

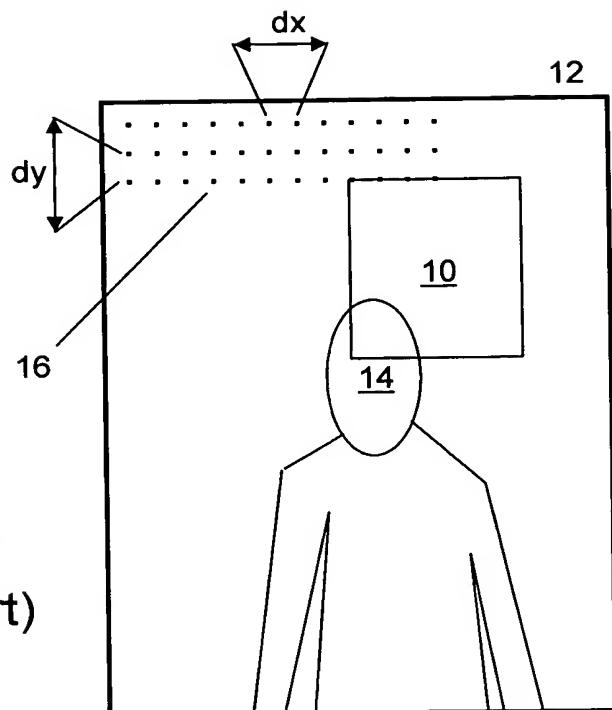


Figure 1  
(Prior art)

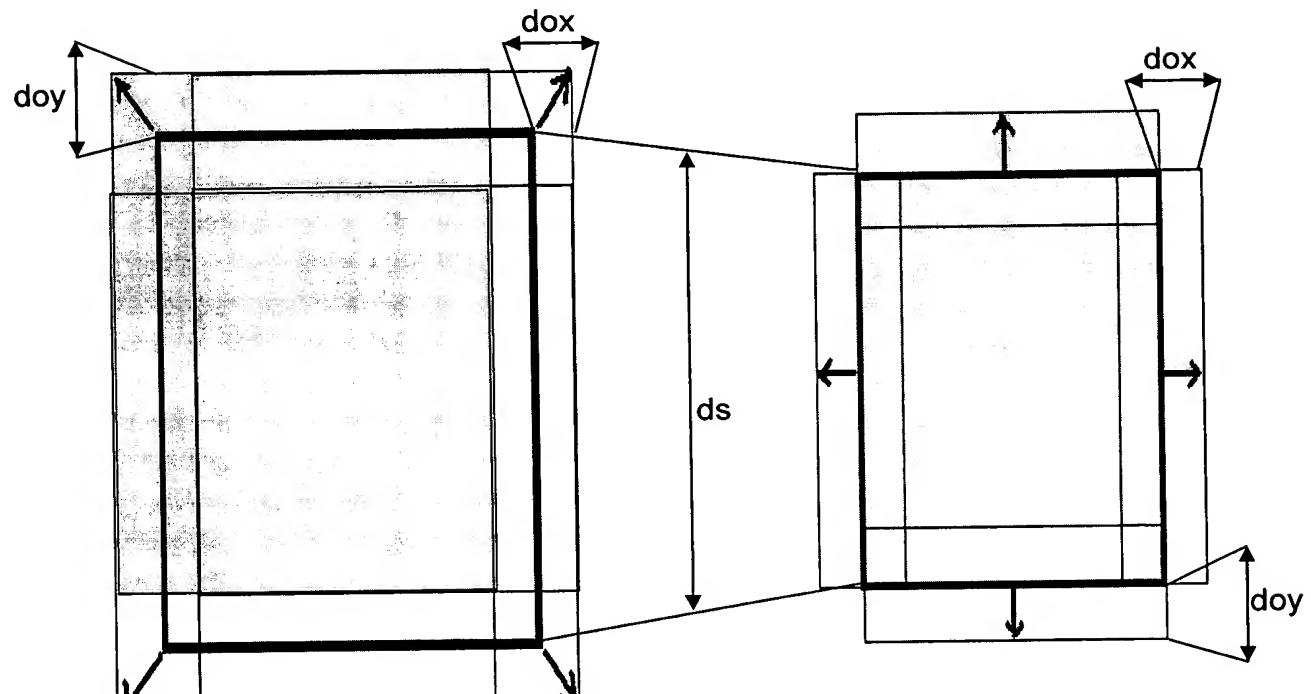


Figure 2(a)

Figure 2(b)

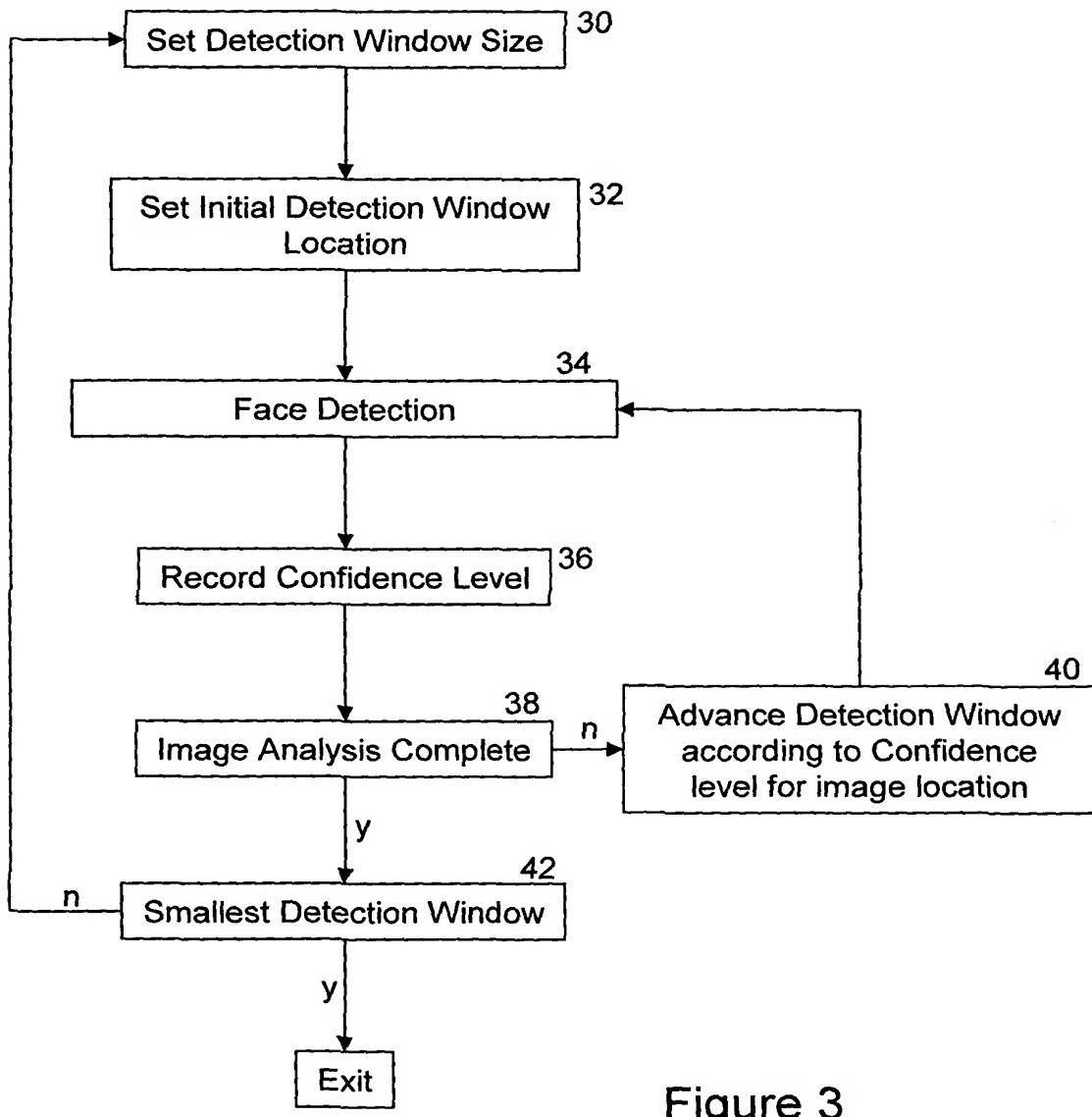
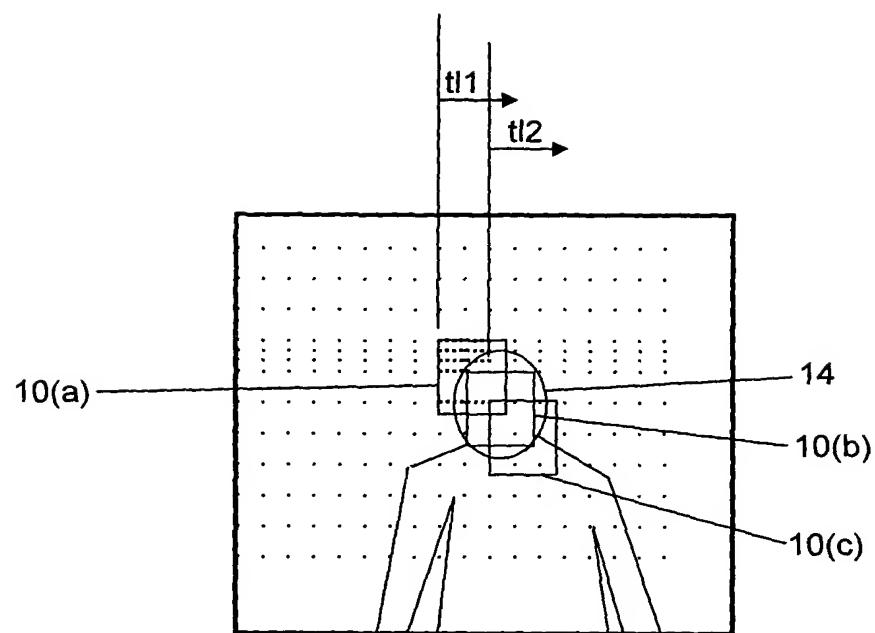
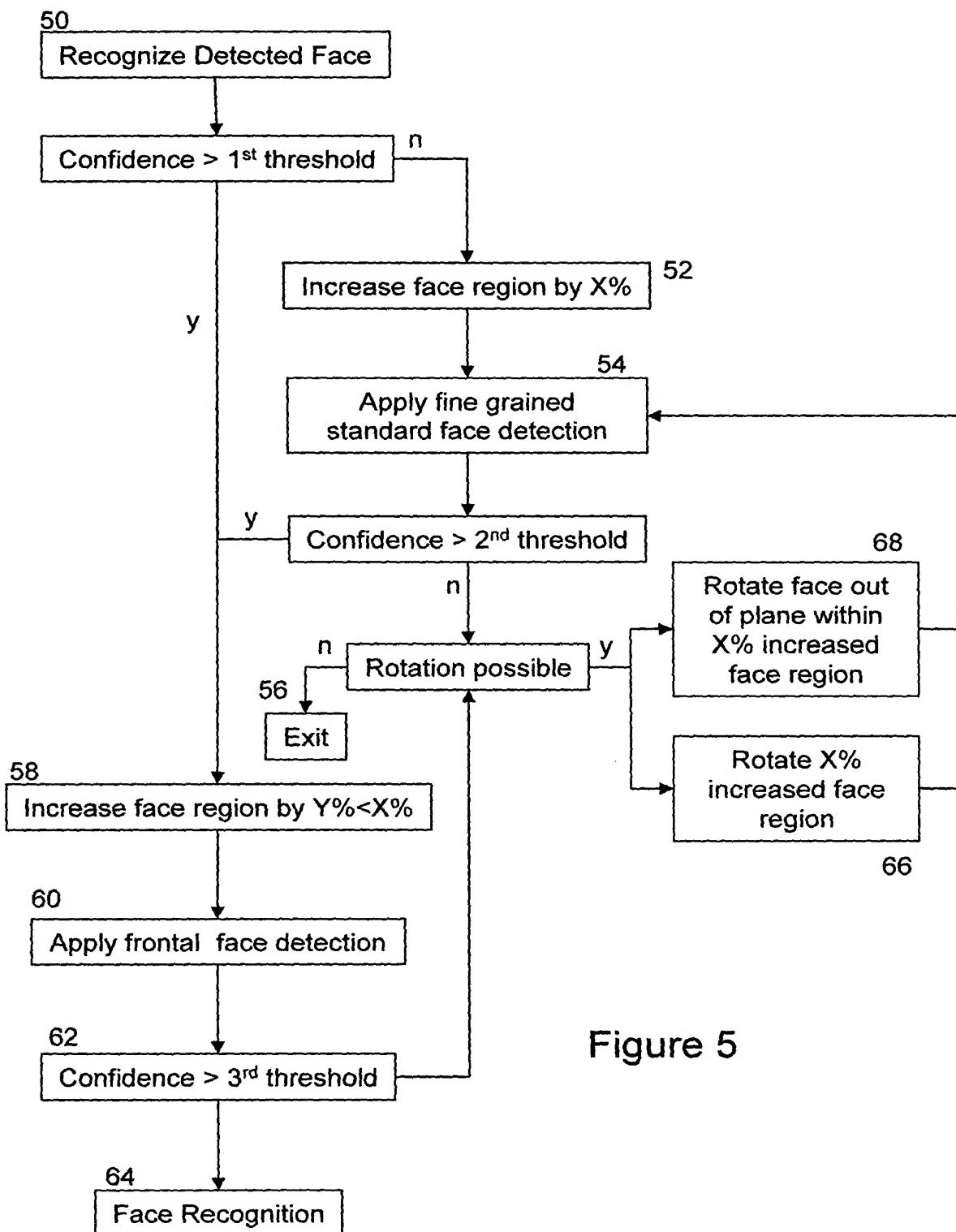


Figure 3



**Figure 4**



## INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2007/006540

A. CLASSIFICATION OF SUBJECT MATTER  
INV. G06K9/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004/264744 A1 (ZHANG LEI [CN] ET AL) 30 December 2004 (2004-12-30)  abstract; figures 3-6 paragraphs [0005], [0061] - [0071], [0076] - [0078], [0082] - [0087] -----	1-16, 18-27, 31,32
X	HENRY A. ROWLEY: "Neural Network-Based Face Detection" May 1999 (1999-05), CARNEGIE MELON UNIVERSITY, PITTSBURG, PA, USA , XP002456221 Chapter 6 page 92 subsection 8.3 ----- -/-	1-9, 11, 19-32

 Further documents are listed in the continuation of Box C. See patent family annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

25 October 2007

08/11/2007

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Alecu, Teodor Iulian

## INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2007/006540

## C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HANJIN RYU ET AL: "Coarse-to-Fine Classification for Image-Based Face Detection" IMAGE AND VIDEO RETRIEVAL LECTURE NOTES IN COMPUTER SCIENCE; LNCS, SPRINGER-VERLAG, BE, vol. 4071, July 2006 (2006-07), pages 291-299, XP019036040 ISBN: 3-540-36018-2 Sections 1 and 3 -----	1-4, 6-28, 31, 32
X	EP 1 626 569 A (FUJI PHOTO FILM CO LTD [JP]) 15 February 2006 (2006-02-15)  paragraphs [0077], [0087] - [0099], [0107] - [0111], [0121] - [0124], [0161] -----	1-9, 11-16, 18, 19, 21, 23-32
X	MEKUZ, N. AND DERPANIS, K.G. AND TSOTSO, J.K.: "ADAPTIVE STEP SIZE WINDOW MATCHING FOR DETECTION" PROCEEDINGS OF THE 18TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, vol. 2, August 2006 (2006-08), pages 259-262, XP002455696 Hong Kong, China Sections 1 and 2 -----	1-6, 11-18, 23-27, 31, 32
X	SMERALDI F ET AL: "Facial feature detection by saccadic exploration of the Gabor decomposition" IMAGE PROCESSING, 1998. ICIP 98. PROCEEDINGS. 1998 INTERNATIONAL CONFERENCE ON CHICAGO, IL, USA 4-7 OCT. 1998, LOS ALAMITOS, CA, USA, IEEE COMPUT. SOC, US, vol. 3, 4 October 1998 (1998-10-04), pages 163-167, XP010586874 ISBN: 0-8186-8821-1 Sections 4 to 6 Sub-section 7.2 -----	1-4, 11, 12, 14-16, 18-20, 22-28, 31, 32
X	FROBA, B., ERNST, A. AND KUBLBECK, C.: "Real time face detection" PROCEEDINGS OF IASTED "SIGNAL AND IMAGE PROCESSING", [Online] 2002, pages 1-6, XP002455698 Kauai, Hawaii Retrieved from the Internet: URL: <a href="http://www.emballi.de/publi/veroeffent/Froeba.pdf">http://www.emballi.de/publi/veroeffent/Froeba.pdf</a> [retrieved on 2007-10-23] Sections 1, 2, 6 -----	1-4, 6-11, 19, 21, 23-32
		-/-

## INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2007/006540

## C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FERAUD R ET AL: "A FAST AND ACCURATE FACE DETECTOR BASED ON NEURAL NETWORKS" IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE SERVICE CENTER, LOS ALAMITOS, CA, US, vol. 23, no. 1, 1 January 2001 (2001-01-01), pages 42-53, XP001008995 ISSN: 0162-8828 Section 2; subsection 2.1 Section 3 -----	1-4, 6-9, 11, 19-28, 31, 32
A	FROBA, B. AND ERNST, A.: "Face detection with the modified census transform" PROCEEDINGS OF THE SIXTH IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE AND GESTURE RECOGNITION, 2004, pages 91-96, XP002455697 cited in the application the whole document -----	1-32
A	US 2006/029265 A1 (KIM JUNGBAE [KR] ET AL) 9 February 2006 (2006-02-09) abstract; figures 1,8 paragraphs [0029] - [0055] -----	1-32

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No

PCT/EP2007/006540

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
US 2004264744	A1	30-12-2004		NONE
EP 1626569	A	15-02-2006	JP 2006285944 A US 2006035259 A1	19-10-2006 16-02-2006
US 2006029265	A1	09-02-2006	KR 20060012777 A	09-02-2006